

CLAIMS

What is claimed is:

- 1 1. A processor-based method comprising:
2 receiving a data stream comprising a plurality of temporally ordered data
3 points;
4 generating a plurality of sequences from a first portion of the data stream; and
5 training a detector by determining a value for a sensitivity parameter using the
6 plurality of sequences.
2. The method, as set forth in claim 1, comprising running the detector on a
second portion of the data stream.
- 1 3. The method, as set forth in claim 2, wherein running the detector comprises:
2 generating a score corresponding to the second portion of the data stream;
3 comparing the score to the determined value for the sensitivity parameter; and
4 signaling detection.
- 1 4. The method, as set forth in claim 1, wherein training the detector by
2 determining the value for the sensitivity parameter comprises selecting the value for the
3 sensitivity parameter based on a target level for an estimated performance characteristic of
4 the detector.

1 5. The method, as set forth in claim 1, wherein training the detector by
2 determining the value for the sensitivity parameter comprises:
3 generating a score for each of the plurality of sequences; and
4 selecting the value for the sensitivity parameter based on the scores.

1 6. The method, as set forth in claim 1, wherein generating the plurality of
2 sequences comprises:
3 inferring a statistical distribution of a known type to characterize the first
4 portion of the data stream; and
5 generating the plurality of sequences from the statistical distribution.

1 7. The method, as set forth in claim 6, wherein the statistical distribution is a
2 discrete distribution containing data points from the first portion of the data stream, and
3 wherein generating the plurality of sequences from the statistical distribution comprises
4 selecting data points from the discrete distribution.

1 8. The method, as set forth in claim 6, wherein inferring a known type of
2 distribution comprises determining a set of parameters corresponding to the known type of
3 statistical distribution.

1 9. The method, as set forth in claim 1, wherein generating the plurality of
2 sequences comprises:
3 selecting a change based on a distribution of changes; and
4 generating a changed sequence based on the selected change.

1 10. The method, as set forth in claim 1, wherein determining the value of the
2 sensitivity parameter comprises determining a plurality of values for the sensitivity parameter
3 using the plurality of sequences.

1 11. The method, as set forth in claim 10, wherein determining one of the plurality
2 of values for the sensitivity parameter comprises calculating a transformation of a second of
3 the plurality of values for the sensitivity parameter.

1 12. The method, as set forth in claim 1,
2 wherein receiving a data stream comprises receiving a plurality of data
3 streams;
4 wherein generating the sequences comprises generating a respective plurality
5 of sequences from a respective first portion of each of the plurality of
6 data streams; and
7 wherein determining the sensitivity parameter comprises determining a
8 respective sensitivity parameter for each of the plurality of sequences.

1 13. The method, as set forth in claim 1, wherein determining the value for the
2 sensitivity parameter comprises determining the value for the sensitivity parameter based at
3 least partially on cost parameters.

1 14. The method, as set forth in claim 12, comprising raising an alarm when a
2 respective detector signals detection when parameterized by the respective sensitivity
3 parameter and run on a respective second portion of a sufficient set of data streams.

1 15. A processor-based method comprising:
 2 training a detector using a plurality of sequences generated from a first portion
 3 of a data stream, wherein the detector is configured to detect an
 4 interesting event in the data stream; and
 5 testing a second portion of the data stream using the trained detector.

1 16. The method, as set forth in claim 15, comprising transforming the data stream
 2 before training the detector.

1 17. The method, as set forth in claim 15, comprising:
 2 generating a plurality of sequences from a third portion of the data stream;
 3 and
 4 retraining the detector using the plurality of sequences generated from the
 5 third portion of the data stream.

1 18. The method, as set forth in claim 15, wherein training the detector comprises
 2 determining one or more sensitivity parameters from the plurality of sequences.

1 19. The method, as set forth in claim 18, wherein testing the second portion of the
 2 data stream comprises:
 3 generating a score associated with the second portion of the data stream; and
 4 comparing the score with the one or more sensitivity parameters.

1 20. The method, as set forth in claim 15, comprising raising an alarm only if an
2 interesting event is detected in the data stream a predetermined number of times within a
3 predetermined amount of time.

1 21. The method, as set forth in claim 15, comprising raising an alarm if the
2 detector detects an interesting event in the data stream.

1 22. A system comprising:
2 a trainer configured to generate a plurality of sequences from a first portion of
3 a data stream and further configured to determine one or more
4 sensitivity parameters based on the sequences; and
5 a detector configured to detect an interesting event in the data stream using the
6 one or more sensitivity parameters.

1 23. The system, as set forth in claim 22, comprising an alarm coupled to the
2 detector and configured to engage when an interesting event in the data stream is detected.

1 24. The system, as set forth in claim 22, comprising an input device coupled to the
2 trainer, wherein the device is configured to allow a user to set cost parameters for use in
3 determining the one or more sensitivity parameters.

1 25. A computer-readable medium storing computer instructions for:
2 generating a plurality of sequences from a first portion of a data stream;
3 determining a sensitivity parameter using the plurality of sequences; and

4 training a detector to detect an interesting event in the data stream using the
5 sensitivity parameter.

1 26. The computer-readable medium, as set forth in claim 25, storing computer
2 instructions for:

3 generating a score corresponding to a second portion of the data stream; and
4 signaling detection of an interesting event in the data stream if the score
5 crosses the sensitivity parameter.

1 27. A system comprising:

2 means for generating a plurality of sequences from a first portion of a data
3 stream;

4 means for determining a sensitivity parameter based on the plurality of
5 sequences; and

6 means for detecting an interesting event in a second portion of the data stream
7 using the sensitivity parameter.

1 28. The system, as set forth in claim 27, wherein means for determining comprises
2 means for generating a plurality of interesting sequences from the data stream, wherein the
3 interesting sequences have a different statistical distribution than a statistical distribution of
4 the first portion of the data stream.

1 29. The system, as set forth in claim 27, wherein means for detecting comprises
2 means for detecting an interesting event in a parameter of the plurality of distributions.

1 30. The system, as set forth in claim 27, comprising means for injecting a change
2 into the first portion of the data steam.